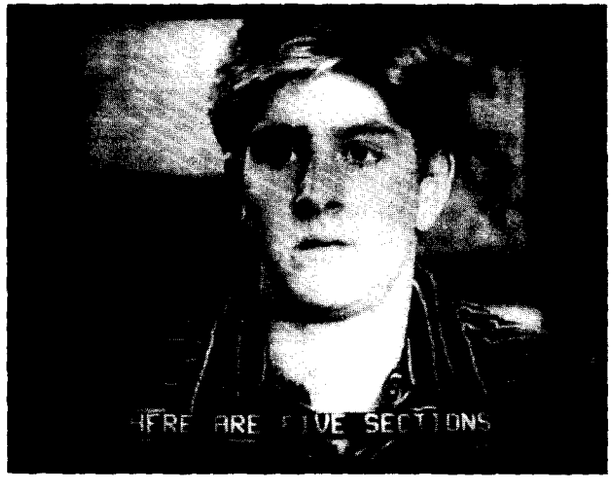
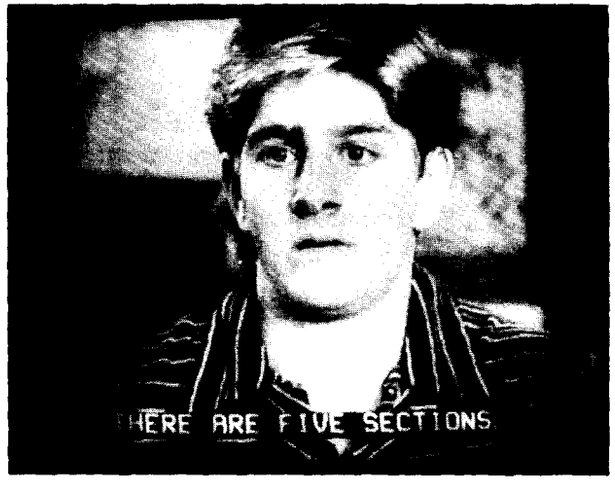
SPEAKER IDENTIFICATION
OFF-SCREEN NARRATOR

In this clip, there are two speakers. One is a narrator. The other speaker is not always visible.





Features used to indicate speakers:

Speaker identification
Adult Kevin:

Capital letters and italics
SLANTED LETTERS/CAPITAL LETTERS

Color
Narrator in white, second speaker in green

Feature	**Preferred**		**Unacceptable**	
	N	%	N	%
Speaker ID	122	65	1	1
Caps & italics	31	18	15	8
Color	36	19	36	19
Totals	189	101*	52	28

*Discrepancies of 1% are a result of standard rounding error.

SPEAKER IDENTIFICATION
MULTIPLE SPEAKERS

In this clip, there are multiple speakers, including more than one off-screen speaker.

Features used to indicate speaker:

Speaker identification
John:

Color
White, green and yellow used for the three speakers

Double chevrons
>>BEFORE EACH PERSON'S LINES

Feature	Preferred		Unacceptable	
	N	%	N	%
Speaker ID	105	56	3	2
Color	59	31	30	16
Double chevrons	25	13	20	11
Totals	189	100	53	29

SPEAKER IDENTIFICATION
MULTIPLE SPEAKERS

ers are on-screen but hard to identify.

Features used to indicate speaker:

Placement and speaker identification
CAPTION IS NEAR THE PERSON SPEAKING and Speaker identified:

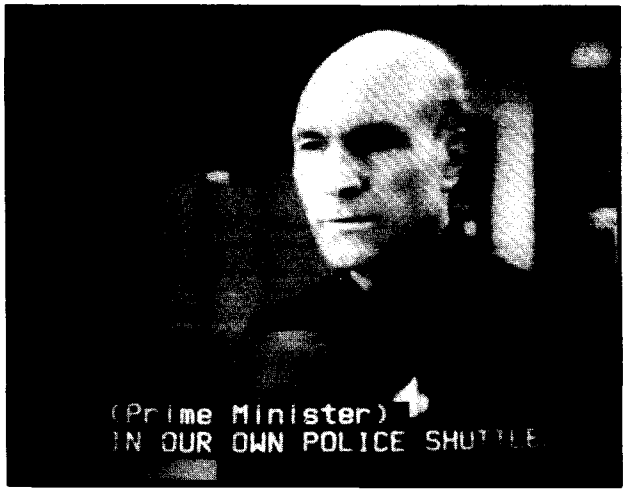
Placement and color
CAPTION IS NEAR THE PERSON SPEAKING and blue, white and yellow used

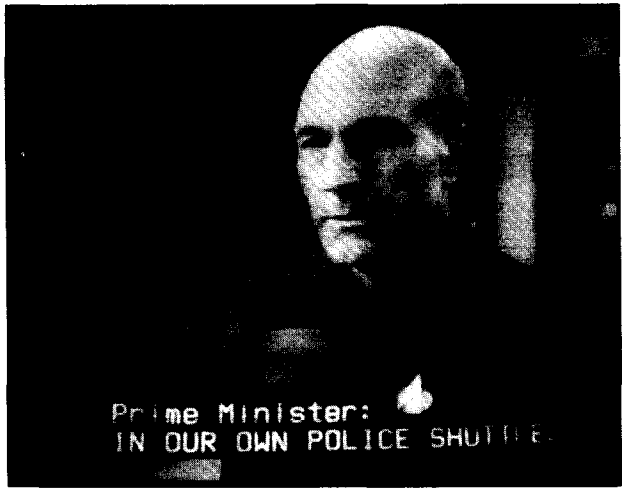
Placement only
CAPTION IS NEAR THE PERSON SPEAKING

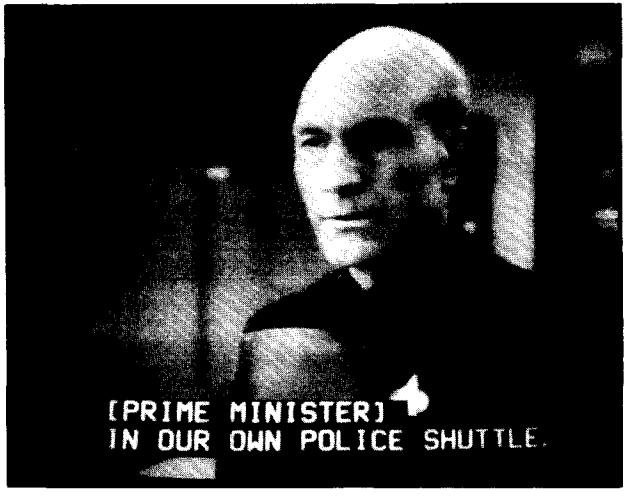
Feature	Preferred		Unacceptable	
	N	%	N	%
Placement & speaker ID	120	64	0	0
Placement & color	44	23	29	15
Placement only	25	13	15	8
Totals	189	100	44	23

SPEAKER IDENTIFICATION
MULTIPLE SPEAKERS

In this clip, there are multiple off-screen speakers.







Features used to indicate speaker:

Parenthesis
(Commander Riker)

Colon
Commander Riker:

Brackets
[COMMANDER RIKER]

Feature	Preferred		Unacceptable	
	N	%	N	%
Parenthesis	77	41	3	2
Colon	57	30	5	3
Brackets	55	29	11	6
Totals	189	100	19	11

SPEAKER IDENTIFICATION

In this clip, a taped phone conversation is played.

Features used to indicate speaker:

Placement, speaker ID, all capital letters
Captions for operator are left; captions for woman are right
911 Operator: EMERGENCY 911 Woman: I NEED AN AMBULANCE

Placement, speaker ID, capital letters and upper/lower case letters
All captions are left
911 Operator: EMERGENCY 911
Woman: I need ambulance

Placement, capital letters and upper/lower case letters
Captions for operator are left; captions for woman are right
EMERGENCY 911I need an ambulance

Feature	Preferred		Unacceptable	
	N	%	N	%
Placement, speaker ID, all caps	78	41	7	4
Placement, speaker ID, caps & upper/lower	57	30	6	3
Placement, caps & upper/lower	54	29	9	5
Totals	189	100	22	12

Sound Effects

In a segment with several sound effects, (1) description, (2) onomatopoeia, and (3) a combination of the two were tested. In this clip, the sound effects were spaced in time so that they did not overlap. A combination of description and onomatopoeia was the preference of more consumers (56%) than was description alone (31%) or onomatopoeia alone (13%). ($\chi^2 = 53.746$, $df = 2$, $p < .001$)

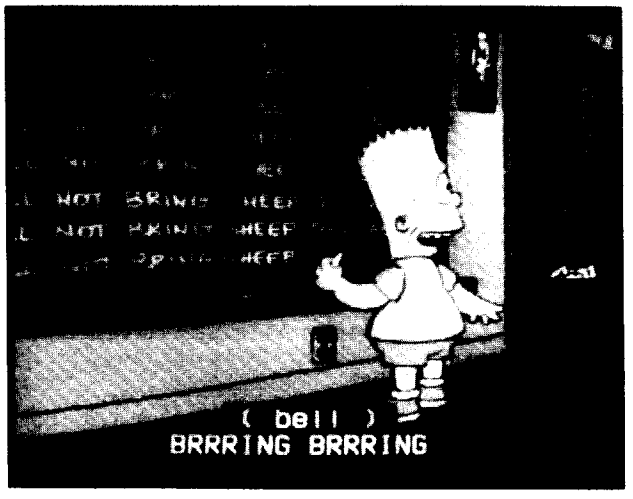
Recommendation: Where feasible, a combination of description and onomatopoeia should be used to indicate sound effects. If space or other limitations do not permit the two to be used together, descriptors should be used. Onomatopoeia should not be used alone. A descriptor is particularly important if the source of the sound effect is not obvious from the video.

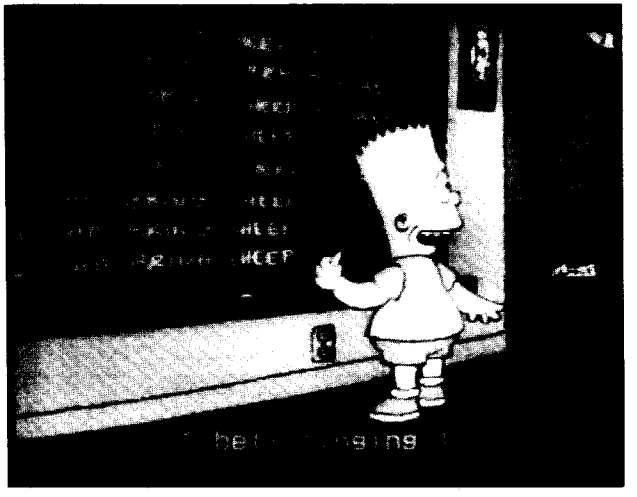
Example:

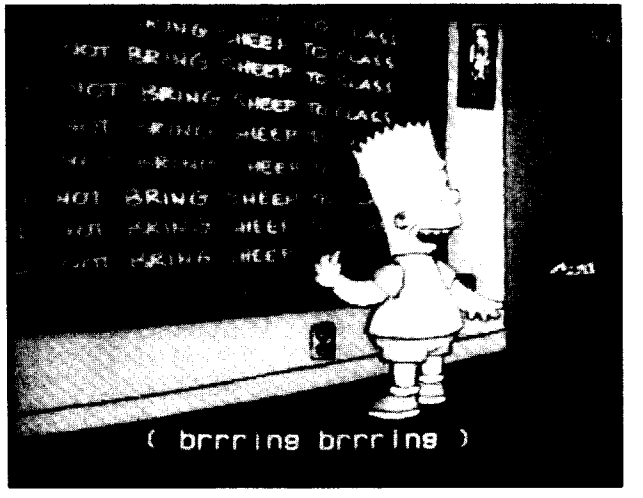
(factory whistle)
TOOOOT

SOUND EFFECTS

In this clip, there are many sound effects.







Features used to indicate sound effects:

Description & onomatopoeia
(whistle)
TOOOOOOOOT

Description
(whistle blowing)

Onomatopoeia
(tooooooooot)

Feature	Preferred		Unacceptable	
	N	%	N	%
Description & onomatopoeia	106	56	6	3
Description	59	31	4	2
Onomatopoeia	24	13	19	10
Totals	189	100	29	15

Music

Background music

In a sentimental scene, (1) a musical-note icon in the upper right corner (to indicate when background music was playing) was tested against (2) explicit description of the background music and against (3) no indication of background music. Description was the rather strong preference (68%) of consumers, with a small minority favoring the music icon (18%) or no indication of background music (15%). ($\chi^2 = 100.793$, $df = 2$, $p < .001$) The music icon and absence of features were also relatively often deemed unacceptable by viewers. In a scene with suspenseful music, 83% preferred description over an absence of any indication of the background music. ($\chi^2 = 80.069$, $df = 1$, $p < .001$)

Recommendation: Background music should be indicated, especially if it contributes to the plot or mood of the video. A description of the background music should be given wherever possible.

Example:

(soft, sad background music)

Singing

The industry convention has been to indicate singing with musical-note icons at the beginning and end of the caption. This style was tested two ways--(1) all caps versus (2) upper/lower case--and also compared with (3) a paint-on style. Some viewers commented that they did not notice the difference between all caps and upper/lower case letters. The paint-on style (a new caption feature) was rejected in this situation, and the convention of music icons combined with captions was validated. In all, 82% of respondents chose styles with the music icons. Of these, slightly more than half chose all caps, and slightly fewer than half chose upper/lower case. The differences among the three choices were significantly significant. ($\chi^2 = 21.555$, $df = 2$, $p < .001$)

Recommendation: Continue the practice of using the musical-note icon surrounding the caption. All-caps and upper/lower-case type are equally acceptable for the caption portion.

Example: ♪ I'VE GOT A SONG TO SING ♪

MUSIC

In this clip, the background music helps to set the mood.

Features used to indicate background music:

Description
(soft, sad background music)

Music icon
music icon appears in top right corner of screen

No features used
music isn't identified

Feature	Preferred		Unacceptable	
	N	%	N	%
Description	128	68	2	1
Music icon	33	18	30	16
No features used	28	15	37	20
Totals	189	101*	69	37

*Discrepancies of 1% are a result of standard rounding error.

MUSIC

In this clip, the background music alerts the viewer that something is going to happen.

Features used to indicate background music:

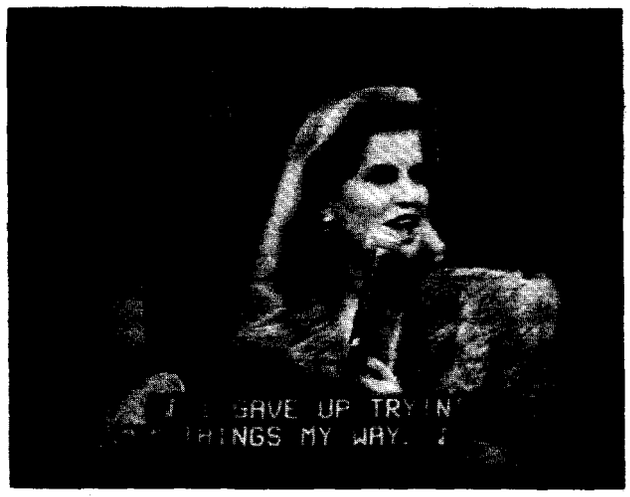
Description
(scary music)

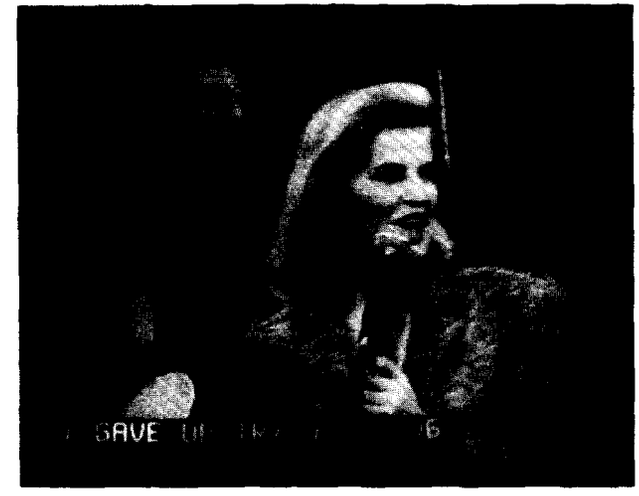
No features used
Music isn't described

Feature	Preferred		Unacceptable	
	N	%	N	%
Description	157	83	2	1
No features used	32	17	27	14
Totals	189	100	29	15

MUSIC

In this clip, a woman is singing.





Features used to indicate singing:

Music icon and capital letters
♪ MUSICAL NOTE AT BEGINNING AND END OF EACH CAPTION BLOCK ♪

Music icon and upper/lower case letters
♪ Musical note at beginning and end of each caption block ♪

Paint-on lyrics
Words appear as she sings

Feature	Preferred		Unacceptable	
	N	%	N	%
Music icon & caps	80	42	2	1
Music icon & upper/lower	76	40	3	2
Paint-on lyrics	33	18	46	24
Totals	189	100	51	27

Audience Reaction

Reaction other than laughter

In a comedy scene, the live audience reacted to an actor's line. Their reaction was captioned in three ways: (1) with a flashing caption (a new feature), (2) with a standard caption, and (3) with a caption combined with identification of the source. Consumers tended strongly (77%) to prefer the descriptive version, where the audience was explicitly identified, compared to flashing captions (13%) or absence of identification of the source of the caption (11%) . ($\chi^2 = 90.063$, $df = 2$, $p < .001$)

Laughter

In a 54-second sitcom clip, the repeated laughter that characterizes sitcoms was treated with (1) a conventional descriptive caption [i.e., (audience laughing)] versus (2) a flashing caption of the words "ha ha ha," versus (3) the absence of any indication of audience laughter. Most consumers (58%) chose description as the best of these three alternatives, and few (5%) found this unacceptable. In the preferred version, the descriptor appeared seven times in 54 seconds. Only 7% preferred flashing captions, and 35% preferred no indication. ($\chi^2 = 71.841$, $df = 2$, $p < .001$)

Recommendation: Audience reaction should be captioned. This is particularly important where the reaction itself becomes part of the plot or comedy. Audience laughter should also be described. (It is of course possible that repeating the descriptor every time the audience laughs, over the length of an entire sitcom episode, would become annoying. This length of exposure was not tested. Therefore, discretion is advised; but audience laughter should be indicated much more often than is now the industry's practice.)

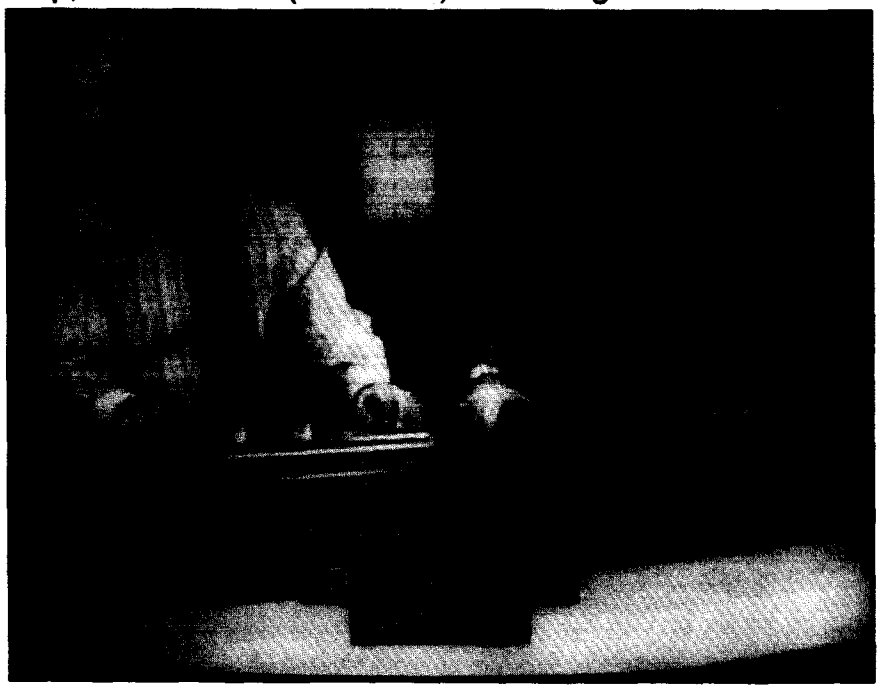
Examples:

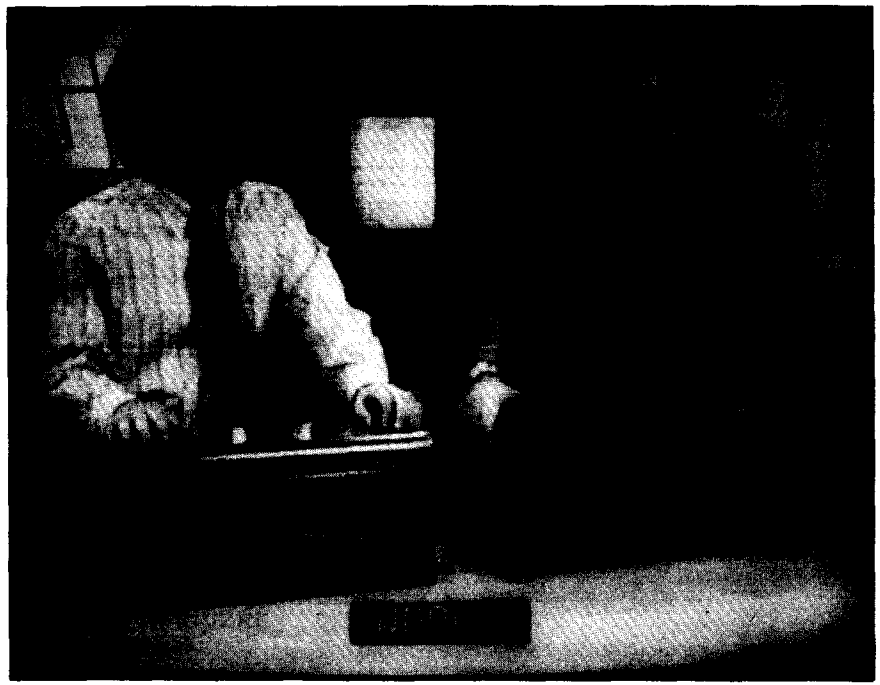
(Audience)
BOO

(audience laughing)

AUDIENCE REACTION

In this clip, the audience (off-screen) is reacting to what the actor has said.





Features used to indicate audience reaction:

Identification
Audience:
WOOH

Flashing
WOOH flashes

No features used
WOOH

Feature	Preferred		Unacceptable	
	N	%	N	%
Identification	145	77	2	1
Flashing	24	13	32	17
No features used	20	11	8	4
Totals	189	101*	42	22

*Discrepancies of 1% are a result of standard rounding error.

AUDIENCE REACTION

In this clip, the studio audience frequently laughs.

Features used to indicate audience laughing:

Description
(audience laughing)

No features used
No mention of audience laughing

Flashing
(ha ha ha) flashes

Feature	Preferred		Unacceptable	
	N	%	N	%
Description	109	58	9	5
No features used	66	35	25	13
Flashing	14	7	53	28
Totals	189	100	87	46

Manner of Speaking

Emotion

In one comedy clip, an angry exchange was captioned with (1) description of the emotion, (2) exaggerated punctuation, and (3) absence of any indication of emotion in the speakers. A plurality of consumers (46%) preferred description of emotion over absence of features (preferred by 28%) and over exaggerated punctuation (preferred by 27%). ($\chi^2 = 13.746$, $df = 2$, $p < .01$) Few (5%) consumers found description unacceptable. These results were statistically significant, but were more mixed than some of the other preference data.

Recommendation: Where strong emotion is being conveyed, the emotion should be described with the caption. This feature should be used especially where the strong emotion is not entirely obvious in the facial expression and actions of the speaker. Caption writers may be concerned that this feature could be over-used. However, based on consumers' reaction, caption writers should use this feature more than is current practice.

Example:

(angrily) DAN? DAN, ARE YOU EVEN LISTENING?

Whispering

Tested were: (1) upper/lower case letters for the whispered lines, but no description; (2) description combined with all capitals; and (3) description combined with upper/lower case captions. (The description in this case was the word "whispering.") Description combined with upper/lower case was the rather strong (70%) preference of consumers, whereas upper/lower case alone was preferred by only 9% and description plus all capitals was preferred by 22%. ($\chi^2 = 115.460$, $df = 2$, $p < .001$)

Recommendation: Whispered lines should be identified as such, as follows, and combined with upper/lower case captions:

Example:

(whispering)
It's dark in here. I can't see you.

Emphasis

When a word or words are emphasized in a line, description is not possible. Three features were tested in a scene in which several significant words were emphasized. The words were indicated with (1) italic captions (2) color (a new feature), and (3) underlining. The reactions to these choices were more mixed than in many of the other items tested. The highest percentage of respondents chose italics (43%), followed rather closely by color (37%). However, color was unacceptable to a larger minority (13%) than was italics (3%). Differences in preference were statistically significant. ($\chi^2 = 17.23$, $df = 2$, $p < .001$)

Recommendation: Indicate emphasized word(s) within a caption with italics.

Example:

DON'T *EVER, EVER* SAY THAT AGAIN.

Accents

A character's southern accent was indicated (1) through description, (2) through exaggerated spelling to indicate the accent, and (3) without indication of accent. The rather strong preference (68%) of consumers was the description without an attempt to otherwise portray the accent. This method was also not controversial, with only one person finding it unacceptable. ($\chi^2 = 103.841$, $df = 2$, $p < .001$)

Recommendation: Indicate foreign or regional accent with a description at the beginning of the character's lines. (Note: This issue was tested only with a fictional character, and probably should not be generalized to other situations.)

Example:

(southern accent)
I'M SURE GLAD I DIDN'T HAVE THAT CASE.

MANNER OF SPEAKING
EMOTION

In this clip, the speakers are angry.

Features used to indicate emotion:

Description
(angrily) DAN? SAY, DAN, ARE YOU EVEN LISTENING TO ME?

No features used
DAN? SAY, DAN, ARE YOU EVEN LISTENING TO ME?

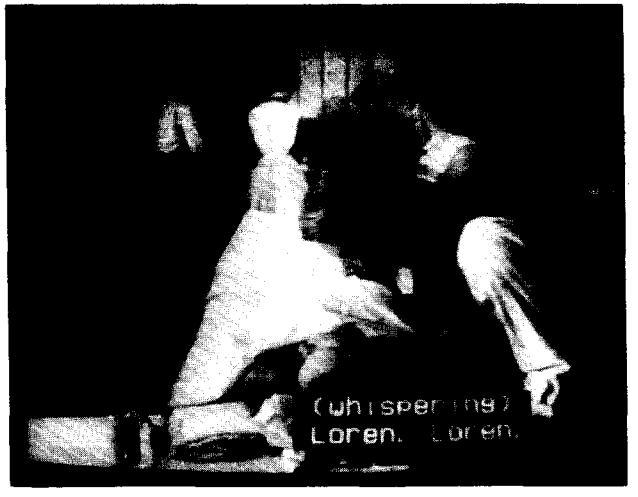
Additional punctuation
DAN?!!!! SAY, DAN, ARE YOU EVEN LISTENING TO ME!!!

Feature	Preferred		Unacceptable	
	N	%	N	%
Description	87	46	10	5
No features used	52	28	5	3
Additional punctuation	50	27	26	14
Totals	189	101*	41	22

*Discrepancies of 1% are a result of standard rounding error.

MANNER OF SPEAKING
WHISPERING

In this clip, one of the actors is whispering.



Features used for indicating whispering:

Description and caption in upper/lower case letters
(whispering)
Loren, Loren

Description and caption in all capital letters
(WHISPERING)
LOREN, LOREN

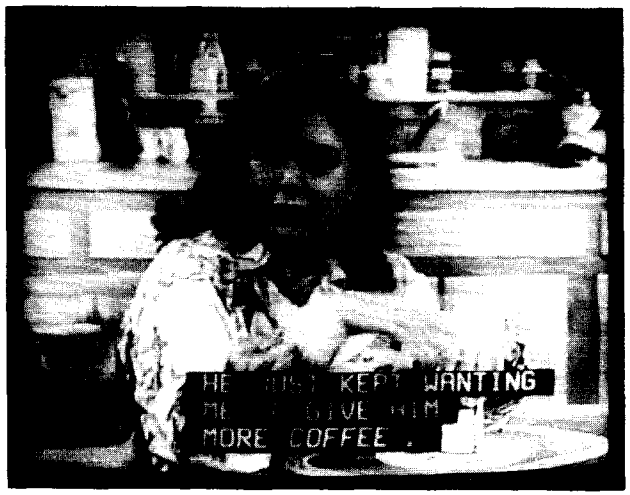
No description and caption in upper/lower case letters
Loren, Loren

Feature	**Preferred**		**Unacceptable**	
	N	%	N	%
Description & upper/lower	131	70	1	1
Description & caps	42	22	15	8
No description & upper/lower	16	9	16	9
Totals	189	101*	32	18

*Discrepancies of 1% are a result of standard rounding error.

MANNER OF SPEAKING
EMPHASIS

In this clip, two captions contain words emphasized by the speaker.



Features used to indicate emphasis:

Italics
COFFEE

Color
COFFEE captioned in green

Underline
COFFEE

Feature	Preferred		Unacceptable	
	N	%	N	%
Italics	82	43	6	3
Color	70	37	25	13
Underline	37	20	21	11
Totals	189	100	52	27

MANNER OF SPEAKING
ACCENT

In this clip, the person speaking has a Southern accent.

Features used for indicating accent:

Description
(Southern accent)

Phonetic transcription
A RACCOOON WITH A FIIISH

No features used
No mention of accent

Feature	Preferred		Unacceptable	
	N	%	N	%
Description	129	68	1	1
Phonetic transcription	32	17	37	20
No features used	28	15	19	10
Totals	189	100	57	31

Explanation of a Pun

Puns can be difficult for deaf and hard of hearing viewers, since enjoyment of the pun tends to depend heavily on hearing. A caption (1) carrying an explanation of the pun below it in upper/lower case was tested versus (2) an absence of such explanation. The explained version was the rather strong preference, with 71% of consumers preferring it and only 1% objecting; 29% preferred a lack of explanation. ($\chi^2 = 32.196$, $df = 1$, $p < .001$)

Recommendation: Puns should be described when feasible.

Example:

I'M LOOKING FOR AMANDA HUGGENKISS
("a man to hug and kiss")

Indication of a Title

Titles have been indicated in various ways by the industry. Quotation marks, italics, and underlining were compared. Quotation marks were favored by 44%, and found unacceptable by only 1%. Italics came in second, with 29%, and underlining third, with 27%. ($\chi^2 = 10.698$, $df = 2$, $p < .01$) Underlining was unacceptable to 14%.

Recommendation: Use quotation marks when indicating the title of a book, movie, etc.

Example:

I GOT THIS CERTAIN URGE TO RENT
"HONEY, I SHRUNK THE KIDS."

EXPLANATION OF A PUN

In this clip, the conversation includes a pun.

Features used to indicate the pun:

Explanation of pun
("a man to hug and kiss")

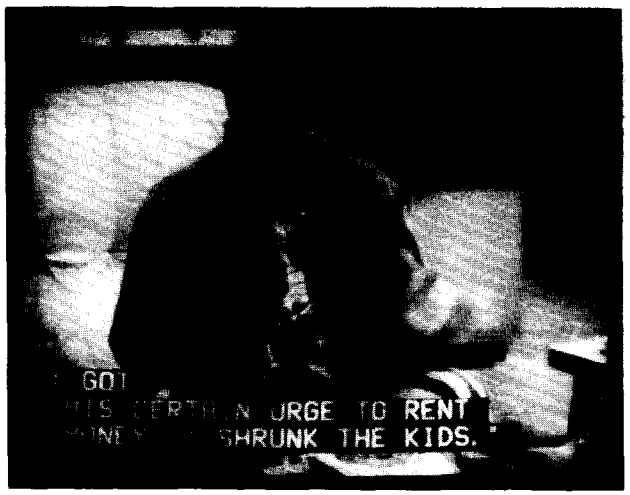
No features used
no explanation

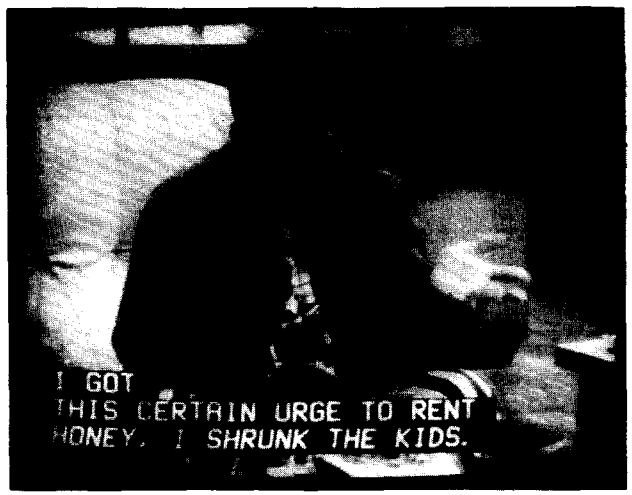
Feature	Preferred		Unacceptable	
	N	%	N	%
Explanation	133	71	2	1
No features used	55	29	5	3
Totals	188*	100	7	4

*One person did not respond to this item.

PART OF LANGUAGE
TITLE

In this clip, the person speaking says the title of a movie.





Features used to indicate movie title:

Quotation marks
"HONEY, I SHRUNK THE KIDS"

Italics
HONEY, I SHRUNK THE KIDS

Underline
HONEY, I SHRUNK THE KIDS

Feature	**Preferred**		**Unacceptable**	
	N	%	N	%
Quotation marks	84	44	2	1
Italics	55	29	3	2
Underline	50	27	26	14
Totals	189	100	31	17

Quantity of Indicators

Two possible constraints to indicating non-speech information are reading rate and complexity of the visual environment. When is there too much information? One set of clips was used as a beginning test of where "too much" may be a factor in consumer preference. A clip with a busy audio background (with sounds carrying information) was captioned at three reading rates: approximately 149 words per minute, 167 words per minute, and 192 words per minute. The difference in reading rate was a factor of the amount of information about background sounds. The fastest, most complex version was the most preferred: 51% preferred this version, 35% preferred the middle version, and only 13% preferred the version with no indications of non-speech information (and a roughly equal number of people found this version unacceptable). ($\chi^2 = 41.523$, $df = 2$, $p < .001$)

Recommendation: If a descriptive caption or feature would in any way clarify or enhance the viewer's experience of being connected with the audio, it should be indicated. Consumers prefer that more of such information be included than is now the case.

AMOUNT OF NON-SPEECH INFORMATION TO INCLUDE

In this clip, different amounts of non-speech information are captioned, in addition to the dialogue.

Features used to indicate amount of non-speech information:

Dialogue and all sounds captioned

Dialogue and some sounds captioned

Dialogue and no sounds captioned

Feature	Preferred		Unacceptable	
	N	%	N	%
All sounds	97	51	8	4
Some sounds	67	35	1	1
No sounds	25	13	24	13
Totals	189	99*	33	18

*Discrepancies of 1% are a result of standard rounding error.